

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: WEI, Ying-Fei
- (ii) TITLE OF INVENTION: Transforming Growth Factor
Alpha HIII
- (iii) NUMBER OF SEQUENCES: 10
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
Cecchi, Stewart & Olstein
 - (B) STREET: 6 Becker Farm Road
 - (C) CITY: Roseland
 - (D) STATE: NJ
 - (E) COUNTRY: USA
 - (F) ZIP: 07068-1739
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: WordPerfect 5.1, Dos Text File
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/778,545
 - (B) FILING DATE: January 3, 1997
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA
 - (A) APPLICATION NUMBER: 60/011,136
 - (B) FILING DATE: January 4, 1996
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: MULLINS, J.G.
 - (B) REGISTRATION NUMBER: 33,073
 - (C) REFERENCE/DOCKET NUMBER: 325800-541 (PF 220)
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 201-994-1700
 - (B) TELEFAX: 201-994-1744

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 923 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GAAA ATG GCG CCT CAC GGC CCG GGT AGT CTT ACG ACC CTG GTG CCC TGG Met Ala Pro His Gly Pro Gly Ser Leu Thr Thr Leu Val Pro Trp -25 -20 -15	49
GCT GCC GCC CTG CTC CTC GCT CTG GGC GTG GAA AGG GCT CTG GCG CTA Ala Ala Ala Leu Leu Leu Ala Leu Gly Val Glu Arg Ala Leu Ala Leu -10 -5 1 5	97
CCC GAG ATA TGC ACC CAA TGT CCA GGG AGC GTG CAA AAT TTG TCA AAA Pro Glu Ile Cys Thr Gln Cys Pro Gly Ser Val Gln Asn Leu Ser Lys 10 15 20	145
GTG GCC TTT TAT TGT AAA ACG ACA CGA GAG CTA ATG CTG CAT GCC CGT Val Ala Phe Tyr Cys Lys Thr Thr Arg Glu Leu Met Leu His Ala Arg 25 30 35	193
TGC TGC CTG AAT CAG AAG GGC ACC ATC TTG GGG CTG GAT CTC CAG AAC Cys Cys Leu Asn Gln Lys Gly Thr Ile Leu Gly Leu Asp Leu Gln Asn 40 45 50	241
TGT TCT CTG GAG GAC CCT GGT CCA AAC TTT CAT CAG GCA CAT ACC ACT Cys Ser Leu Glu Asp Pro Gly Pro Asn Phe His Gln Ala His Thr Thr 55 60 65 70	289
GTC ATC ATA GAC CTG CAA GCA AAC CCC CTC AAA GGT GAC TTG GCC AAC Val Ile Ile Asp Leu Gln Ala Asn Pro Leu Lys Gly Asp Leu Ala Asn 75 80 85	337
ACC TTC CGT GGC TTT ACT CAG CTC CAG ACT CTG ATA CTG CCA CAA CAT Thr Phe Arg Gly Phe Thr Gln Leu Gln Thr Leu Ile Leu Pro Gln His 90 95 100	385
GTC AAC TGT CCT GGA GGA ATT AAT GCC TGG AAT ACT ATC ACC TCT TAT Val Asn Cys Pro Gly Gly Ile Asn Ala Trp Asn Thr Ile Thr Ser Tyr 105 110 115	433
ATA GAC AAC CAA ATC TGT CAA GGG CAA AAG AAC CTT TGC AAT AAC ACT Ile Asp Asn Gln Ile Cys Gln Gly Gln Lys Asn Leu Cys Asn Asn Thr 120 125 130	481
GGG GAC CCA GAA ATG TGT CCT GAG AAT GGA TCT TGT GTA CCT GAT GGT Gly Asp Pro Glu Met Cys Pro Glu Asn Gly Ser Cys Val Pro Asp Gly 135 140 145 150	529
CCA GGT CTT TTG CAG TGT GTT TGT GCT GAT GGT TTC CAT GGA TAC AAG Pro Gly Leu Leu Gln Cys Val Cys Ala Asp Gly Phe His Gly Tyr Lys 155 160 165	577
TGT ATG CGC CAG GGC TCG TTC TCA CTG CTT ATG TTC TTC GGG ATT CTG Cys Met Arg Gln Gly Ser Phe Ser Leu Leu Met Phe Phe Gly Ile Leu 170 175 180	625
GGA GCC ACC ACT CTA TCC GTC TCC ATT CTG CTT TGG GCG ACC CAG CGC Gly Ala Thr Thr Leu Ser Val Ser Ile Leu Leu Trp Ala Thr Gln Arg 185 190 195	673
CGA AAA GCC AAG ACT TCA TGA ACTACAT AGGTCTTACC ATTGACCTAA Arg Lys Ala Lys Thr Ser 200	721
GATCAATCTG AACTATCTTA GCCCAGTCAG GGAGCTCTGC TTCCTAGAAA GGCATCTTTC	781

GCCAGTGGAT TCGCCTCAAG GTTGAGGCCG CCATTGGAAG ATGAAAAATT GCACTCCCTT 841
 GGTGTAGACA AATACCAAGT CCCATTGGTG TTGTTGCCTA TAATAAACAC TTTTCTCTT 901
 TTTAAAAAAA AAAAAAAAAA AA 923

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 229 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ala Pro His Gly Pro Gly Ser Leu Thr Thr Leu Val Pro Trp Ala
 -25 -20 -15 -10
 Ala Ala Leu Leu Leu Ala Leu Gly Val Glu Arg Ala Leu Ala Leu Pro
 -5 1 5
 Glu Ile Cys Thr Gln Cys Pro Gly Ser Val Gln Asn Leu Ser Lys Val
 10 15 20
 Ala Phe Tyr Cys Lys Thr Thr Arg Glu Leu Met Leu His Ala Arg Cys
 25 30 35
 Cys Leu Asn Gln Lys Gly Thr Ile Leu Gly Leu Asp Leu Gln Asn Cys
 40 45 50 55
 Ser Leu Glu Asp Pro Gly Pro Asn Phe His Gln Ala His Thr Thr Val
 60 65 70
 Ile Ile Asp Leu Gln Ala Asn Pro Leu Lys Gly Asp Leu Ala Asn Thr
 75 80 85
 Phe Arg Gly Phe Thr Gln Leu Gln Thr Leu Ile Leu Pro Gln His Val
 90 95 100
 Asn Cys Pro Gly Gly Ile Asn Ala Trp Asn Thr Ile Thr Ser Tyr Ile
 105 110 115
 Asp Asn Gln Ile Cys Gln Gly Gln Lys Asn Leu Cys Asn Asn Thr Gly
 120 125 130 135
 Asp Pro Glu Met Cys Pro Glu Asn Gly Ser Cys Val Pro Asp Gly Pro
 140 145 150
 Gly Leu Leu Gln Cys Val Cys Ala Asp Gly Phe His Gly Tyr Lys Cys
 155 160 165
 Met Arg Gln Gly Ser Phe Ser Leu Leu Met Phe Phe Gly Ile Leu Gly
 170 175 180
 Ala Thr Thr Leu Ser Val Ser Ile Leu Leu Trp Ala Thr Gln Arg Arg
 185 190 195
 Lys Ala Lys Thr Ser
 200

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CGCGGATCCG GGCAAAAGAA CCTTTGC

27

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCGTCTAGAC TAAAGCAGTG AGAACGAGCC

30

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGCGGATCCG TCCATCATGG CGCCTCACGG CCCG

34

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GCGTCTAGAC TACATAAGCA GTGAGAACGA GCC

33

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGCGGATCCC GGGCAAAAGA ACCTTTGC

28

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GCGTCTAGAC TACATAAGCA GTGAGAACGA GCC

33

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCGCTCAGAC ATAAGCAGTG AGAACGAGCC

30

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 52 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Gln	Gly	Leu	Asn	Lys	Cys	Asn	Asn	Thr	Gly	Asp	Pro	Glu	Met	Cys	Pro
1				5					10					15	
Glu	Asn	Gly	Ser	Cys	Val	Pro	Asp	Gly	Pro	Gly	Leu	Leu	Gln	Cys	Val
			20					25					30		
Cys	Ala	Asp	Gly	Phe	His	Gly	Tyr	Lys	Cys	Met	Arg	Gln	Gly	Ser	Phe
		35					40					45			
Ser	Leu	Leu	Met												
	50														